physical component of quality of life. Particularly severe back pain and dizziness in combination with falling, and 5 or more comorbidities should be taken into account in OA patients undergoing joint replacement surgery, in order to tune to expectations, satisfaction and outcome after surgery.

317
THE NEED FOR CAUTION IN THE SELECTION AND INTERPRETATION OF MEASURES OF FUNCTION FOR PATIENTS WITH SEVERE HIP AND KNEE PROBLEMS

E. Lenguerrand †, V. Wylde †, L. Brunton †, R. Gooberman-Hill †, P. Dieppe ‡, A. Blom †. † *Univ. of Bristol, Bristol, United Kingdom;* ‡ *Univ. of Exeter, Exeter, United Kingdom*

Purpose: Joint replacement is a common operation, predominantly provided to older adults with osteoarthritis. Reduced function is one of the main indications for joint replacement, and recovery of function one of its main objectives. Adequate assessment of function is critical to help both health care professionals and patients decide upon treatment options and to enable evaluation of the effectiveness of joint replacement.

Physical function is commonly assessed in three ways: patient-reported outcome measures (PROMs), performance tests, and clinician-administered measures. It is recommended that several types of these measures should be used concurrently to capture an extended picture of function but patient burden, resources and logistical constraints mean that such an approach is seldom feasible. Moreover, most of the tools in use do not clearly differentiate between impairments, activity limitations and participation restrictions, and it is unclear to what extent each measure describes them, which in turn has implications as to their substitutability and comparability. The picture is further complicated by the effects of multi-morbidity and patient characteristics on functional measures.

The aims of this research were twofold: 1) to describe the correlations between a variety of commonly used functional measures collected on the same group of older patients listed for hip or knee replacement, and 2) explore associations between these measures and patient characteristics.

Methods: 1,451 patients listed for primary or revision hip or knee replacement at a single high-volume orthopaedic centre were invited to take part in a study of function before joint replacement. 264 agreed to do so and provided their informed consent.

Participants were asked to complete a postal questionnaire about their age, gender, BMI, living arrangements, education level and working status. They completed the Functional Co-morbidity Index and the Hospital Anxiety and Depression Scale. The severity of arthritis was derived as a count of affected joints other than the joint listed for surgery. Two PROMs were also completed, the Western Ontario McMaster Arthritis Index pain and function scales and the Aberdeen Impairment, Activity Limitation and Participation measure.

Participants were also invited to an appointment during which clinician-administered measures and performance tests were performed. These comprised the Harris Hip Score (HHS) or the American Knee Society Score (AKSS), which were completed by a research nurse. The performance tests were a timed 20 metre walk, the get-up-and-go test, step tests (20 and 30 cm high blocks), and a single stance balance test. The relationships between the functional measures were assessed with Spearman Rank coefficient, point-biserial coefficient or Cramér's V statistic. The association between participants' characteristics and functional measures were investigated with linear regression or modified Poisson regression with robust error variance. Individual patient characteristics were first considered in univariate models and then in multivariate analyses to determine if their effects were confounded by other factors. The analyses were conducted separately for hip and knee patients.

Results: Strong to moderate correlations were found within PROMs ($\rho = 0.63$ to 0.88) and within performance tests, but correlations were weaker between these two assessment approaches ($\rho = 0.17$ to 0.65). The HHS correlated better with PROMs ($\rho \sim 0.70$) than performance tests ($\rho = 0.38$ to 0.67); poor correlations were found with the AKSS and other functional measures ($\rho = 0.18$ to 0.28).

Patients' psychological well-being was associated with PROMs (p-value <0.0001) but not with performance tests. Age was associated with performance tests (p-values ranging from <0.05 to <0.0001) but not with PROMs. Pain was strongly associated with function irrespectively of the measurement method. Other patient characteristics had weak or no association with function.

Conclusions: We found that PROMs, performance tests and clinician-administered measures provide information on different aspects of function. This study also shows that associations between function and patient characteristics differed according to the measurement approach used. The functional limitations caused by joint pathology are entangled with the limitations resulting from ageing when measured with performance tests; while psychological status seems to play an important role in patients' self-reported perception of functional ability. Both objective and subjective measures need to be adjusted for pain to enable appropriate interpretation.

318
CLINIMETRICS OF THE STAIR CLIMB TEST IN THE AMSTERDAM COHORT

J.H. Abbott †, M. van der Esch ‡, L.D. Roorde ‡, J. Knoop ‡, M. van der Leeden ‡, W.F. Lems ‡. † *Univ. of Otago, Dunedin, New Zealand;* ‡ *Amsterdam Rehabilitation Res. Ctr., Reade, Amsterdam, Netherlands*

Purpose: OARSI recently published recommendations for physical performance tests in hip or knee osteoarthritis (OA). A test of stair climbing was recommended, but the authors could not recommend a specific stair climbing test or report clinimetric properties due to a lack of data in the literature. The purpose of this study was to describe clinimetric properties of the stair climb test (SCT) in people with knee OA. SCT-up and SCT-down were reported separately, to reveal potentially differing clinimetric properties.

Methods: Baseline and 2-year follow-up data from the Amsterdam Cohort ($n = 200$) were analyzed. Construct validity (convergent) was estimated using Pearson's correlation coefficients between the SCTs and the Western Ontario and McMaster Osteoarthritis index (WOMAC). Smallest detectable change (SDC) at the 90% confidence level was calculated from the standard error of measurement (SEM). We report the difference in SCT means between responders and non-responders at different cut-points of WOMAC Total scale score (WOMAC-T); WOMAC physical function subscale score (WOMAC-PF); and knee extensor strength change reported in the literature. Minimum important change (MIC) was calculated using receiver operating characteristics (ROC) curve methodology, from baseline and follow-up data. The change data was dichotomised into responders and non-responders using cut-off criteria for each of the following external references: WOMAC-T; WOMAC-PF; and knee extensor strength change. MIC estimates for small, medium and large change are reported for each of the cut-points sourced from the literature. Construct validity (discriminant) was estimated using the area under the ROC curve (AUC), with an a priori significance level of .80.

Results: Construct validity (convergent) was moderate (.39) against WOMAC-T and WOMAC-PF (.41). SDC (90%) was larger for SCT-down (3.75 sec) than for SCT-up (1.63 sec). The proportion of patients achieving MCID on the WOMAC-T was 44.5% and for small change and 33.5% for moderate change. For WOMAC-T, the difference in mean SCT time for those who achieved small change in SCT-up was .82 sec (95%CI .27, 1.4) and .93 sec (.35, 1.5) for moderate change. The difference in SCT-up means between responders and non-responders on the WOMAC-PF were .87 (.27, 1.50), .89 (.08, 1.38) and .74 (.09, 1.38) for small, moderate and large changes, respectively. For knee extensor strength, the differences in SCT-up means were: .46 (−.08, .99), .43 (−.15, 1.01), .36 (−.31, 1.03), and .09 (−.66, .84) for small, moderate, medium and large changes respectively. MIC for the SCT-up was −.40 for small change and −0.70 for moderate change in WOMAC-T. MIC for the SCT-up based on small, moderate and large changes in WOMAC-PF were −.40, −.40 and −.30, respectively. MIC for the SCT-up based on small, moderate, medium and large changes in knee extensor strength was −.70 at all levels. MDC and MIC values for the SCT-down were larger or equal in 7 of 9 analyses. AUC values did not exceed .75 for any ROC analysis.